基于 XGBoost 方法的社交网络异常用户检测技术 *

袁丽欣，顾益军†，赵大鹏

(中国人民公安大学 信息技术与网络安全学院，北京 102600)

摘 要：针对传统的社交网络异常用户检测算法应用于现实中非平衡数据集时存在召回率低、运行效率低等问题，对社交网络数据集提取用户内容、行为、属性、关系特征，应用梯度增强集成分类器 XGBoost 算法进行特征选择，建立分类模型，构造非平衡数据集并识别三类垃圾广告发送账号。实验结果表明，该方法与随机森林等传统分类方法相比，对平衡及非平衡数据集进行异常用户检测均实现召回率和 F_1 值的有效提升；选取少量特征同样可达到较高检测水平，证明了方法的有效性。

关键词：XGBoost；社交网络；异常用户检测；异常账号检测；垃圾广告发送者

中图分类号：TP393 doi: 10.19734/j.issn.1001-3695.2018.08.0651

Research on abnormal user detection technology in social network based on XGBoost method

Yuan Lixin, Gu Yijun†, Zhao Dapeng

(*School of Information Technology & Network Security Enforcement, People's Public Security University of China, Beijing* 102600, *China*)

Abstract: Aiming at the problems of low recall rate and poor running efficiency caused by traditional abnormal accounts detecting algorithms in non-balanced social network datasets, the paper extracted user content, behavior, attributes, and relationship features from social network data sets, selected features using gradient-enhanced ensemble classifier XGBoost algorithm, established classification model, constructed unbalanced data sets and realized the identification of three types of spam accounts. Experimental results that the recall rate and the F1 value in identification of three types of abnormal users are improved effectively by XGBoost algorithm in binary classification and multiple classification tasks both in the balanced and unbalanced dataset in comparison with the traditional classification methods such as random forest. And with few features selected by XGBoost, the classification algorithms can get the same effect as with all features of samples, which proved the effectiveness of the method.

Key words: XGBoost; social networks; abnormal users detection; abnormal accounts detection; spam

0 引言

近年来，社交网络和社会媒体得到蓬勃发展，然而以垃圾广告发送者（又称 spam 账号）为主的异常用户时刻污染着社交网络环境[1]。该类账号是攻击者创建的用于发布广告、钓鱼、色情等 URL 的虚假用户，具有较为明显的行为特征。

它们利用在线社交网络大规模传播有害信息，干扰平台的正常使用，威胁着互联网安全[2]。快速有效地识别 spam 账号有助于从源头上净化社交网络环境，保障互联网安全，是公安舆情领域和学术界的重点研究问题之一。

1 相关工作

1.1 现有检测技术

当前，学术界的社交网络异常用户检测工作普遍是对社交网络的节点提取包括注册属性、发布内容、活动行为、连接关系等在内的一类或几类特征，构建多维特征向量，再运用机器学习等方式进行检测，可划分为基于监督学习和无监督学习的方式。

1.2 无监督学习检测方法

无监督学习检测方法是直接根据待检测样本的多维特征进行聚类，从而将正常用户和 spam 用户聚集为不同的簇的方法。由于该方法不需要训练样本，因此可以快速形成检测系统。Miller 等[3]利用 Twitter 用户个人信息和文本内容特征将正常用户和 Spam 账号聚为不同的类；Chu 等人[5]通过 Twitter 发布内容中嵌入的 URL 的最终跳转地址对微博进行聚类，并判断类内账号是否为 spam 账号。

1.3 监督学习检测方法

监督学习的检测方法利用事先标注类别的数据集训练分类模型，再将模型运用到未标注数据中进行预测。Zheng X[8]、吕少卿[9]等利用账号创建时间、消息评论数等内容和行为特征构建分类器，检测 Spam 账号；刘琛[10]根据用户发布微博频率、博文中“@”个数等行为特征建模并识别过度转发、关注行为及虚假粉丝；Meng Jiang[11]、Xue[12]对社交网络关系图中节点的入度、出度和影响力进行建模，检测关注量与好友数不匹配的虚假账号；F.B 等人采用随机森林和 SVM 方法检测 spam 用户，并公开了数据集[1]。

传统的监督学习分类方法包括支持向量机、随机森林等。其中支持向量机（SVM）通过在高维向量空间寻找超平面来实现样本分类，计算复杂度低、对小样本数据分类效果出众，尤其适用二分类任务。随机森林（random forest，RF）等基于决策树的集成分类模型训练时每次从 n 维原始特征中选择 k 个最有效特征进行分裂（k<n），并行地生成多棵决策

收稿日期：2018-08-17；修回日期：2018-10-16 基金项目：国家重点研发计划资助项目（2017YFC0820100）

作者简介：袁丽欣（1993-），女，山西人，硕士研究生，目前主要研究方向为网络情报技术；顾益军（1968-），男（通信作者），江苏人，副教授，博士，主要研究方向为网络情报技术(guyijun@ppsuc.edu.com)；赵大鹏（1994-）男，河南人，硕士研究生，主要研究方向为网络情报技术.

树投票决定分类结果，对多维特征数据分类具有优秀的检测效果。

1.4 当前检测方法的局限性

由于无监督学习只能将内在特征相似的用户聚集为簇，但无法直接确定簇的分类标签，监督学习方式能有效利用社交网络账号多维度的特征，直接预测分类标签，生成的分类模型准确性更高。因此采用监督学习方式进行异常用户检测更为有效。当前常用的监督学习方法虽然能够达到一定的检测目标，但检测精度依然有限，这主要由特征选择和算法选择两方面的原因引起的。

a)特征选择方面，前人研究往往仅选择行为特征等同一类特征进行检测。通常，由于社交网络异常用户的多类特征均与正常用户有所区别，因此只选择某类特征容易遗漏其他特征所蕴涵的信息，不足以准确描述数据的真实情况，导致检测效果不佳。但如若选择全部特征，由于社交网络账号各特征之间存在相关性，SVM 等运用 embedding 方法将样本非正交的多维特征直接投射为正交向量空间的方式容易造成偏差,对高维特征的检测效果很有限。因此需要寻求一种特征选择方法，达到既利用全类别特征，同时避免高维特征引发噪声的目标。

b)算法选择方面，随机森林方法虽能通过特征选择过程降低数据集维度、消除非正交特征影响，但每一次分裂中未被选中的特征无法参与本轮运算，容易造成特征信息损耗，产生误差。并且，由于社交网络真实数据集是一个正常用户数量远超异常用户的非平衡数据集，存在长尾效应，随机森林在不平衡数据集上检测时会出现分类效果不佳、泛化误差变大等一系列问题,因此需要选择一种能有效利用多维特征、并且样本集严重不均衡时依然有效的算法。

当前，对非平衡数据进行分类是社交网络异常检测的研究难点之一。学术界对非平衡数据分类问题的解决方式主要包括利用重采样技术[15~17]以及改进分类算法[18, 19]。重采样技术通过扩大较小类数据规模或缩小较大类规模的方式降低类间非平衡率，但通过欠采样或过采样构造而来的新数据集无法完全符合原始数据集的真实分布，容易造成信息损耗或过度拟合。通过在原有算法的基础上引入增量联机学习[18]、集成学习[19]等方法进行改进，也可以实现降低算法在非平衡数据集的敏感度，但该方式容易引入计算复杂度大、效率低等新问题，并且仅在单一层面关注解决非平衡分类问题的同时容易牺牲模型的泛化性。

2 基于 XGBoost 的异常用户检测方法

社交网络异常用户检测的本质是将数据集中的所有样本划分为正常用户及各类异常用户的多分类任务。本文选择 XGBoost（extreme gradient boosting）[13]集成提升方法构建分类模型。分类训练数据集的每一个样本对应社交网络中的每一个用户，由包含内容、行为、属性、关系等在内的 n 维特征向量 xi 和对应的 p 个类别标签 yi 构成：$\{x_i, y_i\}_{i\in[1,m]}$，$x_i \in \Re^n$，$y_i \in \{class_1, class_2, \ldots, class_p\}$。基于 XGBoost 对用户进行分类的方法是通过学习输入的训练样本，构造分类模型，挖掘特征取值 xi 与类别标签 yi 的关系 $f(x_i)=y_i$，从而预测新样本的类别。整体检测流程如图 1 所示。

对本文提出的分类任务，XGBoost 每一轮训练都是在上一轮的基础上迭代产生的，第 t 次迭代对生成树构造的目标函数为

$$Obj^{(t)}=\sum_{i=1}^{n} l(y_i,\hat{y}_i^{(t)})+\sum_{i=1}^{t}\Omega(f_i)=\sum_{i=1}^{n} l(y_i,\hat{y}_i^{(t-1)}+f_t(x_i))+\Omega(f_t)+constant \quad (1)$$



图 1　基于 XGBoost 的社交网络异常用户检测流程

Fig.1　Abnormal user detection process in social network based on XGBoost

每轮生成的树模型由结构部分 q 和叶子节点样本权重 w 共同表示为：$f_t(x)$　$w_q(x)$。树的复杂度由叶子个数 T 和样本权重 w 的 L2 模平方共同决定，其中 T 越大，样本间的 w 值越不均匀，则树的结构越复杂。正则化项 (f_i) 控制模型的复杂程度，有效防止过拟合，定义为

$$\Omega(f_i)=\gamma T+\frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \quad (2)$$

将目标函数二阶泰勒展开，改写后得到最终目标函数为

$$Obj^{(t)} \simeq [G_j w_j + \frac{1}{2}(H_j+\lambda)w_j^2]+\gamma T \quad (3)$$

其中：$I_j=\{i \mid q(x_i=j)\}$ 表示各叶子节点上样本集合，$G_i=\sum_{i\subset I_j} g_j, H_j=\sum_{i\subset I_j} h_j$ 分别为叶子节点各样本一阶、二阶导数之和。共同运用一阶、二阶导数信息进行优化可得到整体最优解。

实验通过每一步尝试对已有叶子节点加入分隔来逐渐生成最优的树结构，分裂的增益为

$$Gain=\frac{1}{2}[\frac{G_L^2}{H_L+\lambda}+\frac{G_R^2}{H_R+\lambda}-\frac{(G_L+G_R)^2}{H_L+H_R+\lambda}]-\gamma \quad (4)$$

当分裂增益连续小于定值或分裂次数达到指定的最大深度时停止分裂，得到最终分类模型。

对于上文提到的特征选择问题，本文构建分类模型时保留用户的内容、行为、属性、关系等全部特征，充分利用各类特征的有效信息，避免信息损耗；通过串行迭代运算寻找损失函数最优值来优化树的结构，消除样本非正交特征的影响。初次训练完毕后，利用 XGBoost 统计各特征被用于决策树分裂的次数，计算样本特征与分类结果的关联程度，从而按照特征重要性进行特征选择，降低维度。

面向现实中 spam 用户数量远少于正常用户的不平衡数据，本文进行多次集成迭代运算，并且控制 XGBoost 的 max_delta_step 参数来限制每棵树的权重，改变最大步长，从而避免小数量类别中实例样本对分类结果的影响程度过大，降低训练数据不平衡造成的误差。

3 实验分析研究

3.1 数据集和对比方法

本文采用 Apontador 数据集[1]检验方法的有效性。该数据集由巴西著名的基于位置的社交网络采集而来，是包含正常用户和 spam 用户的平衡数据集,其中 spam 用户包括 3 类，分别是产品营销广告发布者（LM）、发布内容与话题标签信息不符的内容污染者（PL）、以及攻击、谩骂等不良言论发布者（BM），分别占异常用户比例为 31%、48.5%、21.4% 。每条记录包含 59 个特征字段（表 1）和 2 个分类字段。

原作者利用分别使用支持向量机（SVM）和随机森林（RF）方法，对该数据集的 4 类用户进行了①直接分类，②先区分样本是否属于异常、再区分异常用户类别的二次分类，验证了 RF 在以上分类任务中效果均明显优于 SVM（直接分类时 RF 对三类 spam 的召回率比 SVM 分别提升 3.2%，4.5%，5.8%，二次分类中分别提升 1.7%，3.9%，6.3%）。为体现本文方法的合理性，本文在 python 环境中复现文献[1]最优参数下的 RF 分类实验，将其作为实验对比。

表 1　特征说明表

Table 1　Feature list

特征类别	特征名
内容特征(32)	内容中邮箱数、URLs 数、电话号码个数、数字字符数、用户所有 tip 中的 1-gram、2-gram、3-gram 的数量及比例、相似度评分（平均值、中位数、最大最小值、标准差）、垃圾邮件关键词数量、攻击值、大写字母数、攻击性词汇数、“this helped me”点击数、“举报滥用行为”点击数、SASA、大写单词数、幸福指数、PANAS-t、Combined-met hod、SentiStrength、SentiWordNet、SenticNet
用户特征(32)	发表 tip 数、发布照片数、注册地点数、评论地点的总距离（平均值，中位数，最大值，最小值，标准差）、用户的熵、信息覆盖的区域个数、用户关注的主题数
地点特征(5)	发布包含地点的 tip 数、地点评分、“赞”点击数、“踩”点击数、地点主页点击数
关系特征(12)	节点出入度比、节点度/邻居节点平均度、聚类系数、粉丝数、关注数、pagerank 值、双向关注比、节点相关性、节点中心性

3.2　实验步骤及参数选择

本实验在 macOS 10.13.4 系统、2.9 GHz Inter Core i5 处理器、Python 3.6.4 环境下进行，步骤如下：

a)读入数据并进行预处理。读入数据，检查数据的格式和分布，利用 XGBoost 计算特征影响力排名，进行特征重要性排序。

b)将原始数据集采用 5 折交叉验证方法划分实验集、测试集，循环评估模型分类效果。划分时对各个类别的样本进行随机分层取样，确保训练集、测试集中各类样本的分布与原始数据集相同，避免采样误差。

c)训练 XGBoost 模型并调参。对 b)得到的每一组训练集采用 5 折交叉验证的方式划分训练子集、验证子集，在训练子集中利用 XGBoost 方法迭代训练模型，运用 CV 网格搜索的方式分别选取各个参数的最优值，逐步调参，并利用验证子集验证模型分类效果，选取最优参数组。

d)选择由最优参数训练而成的模型在测试集中预测分类结果，输出混淆矩阵，计算准确率 P、召回率 R、F1 值等评价指标。

经验证，参数为 max_depth=3，n_estimators=100，n_threthould =None 时，XGBoost 可获得最优分类效果。如图 2、3 所示。

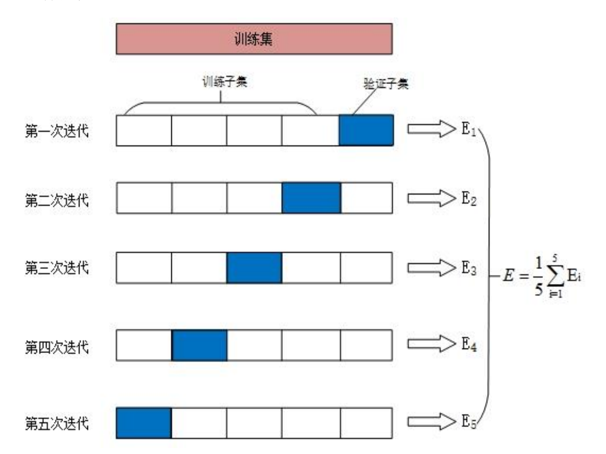


图 2 交叉验证过程展示图

Fig. 2　Cross validation process

3.3　实验结果分析

3.3.1 平衡数据集检测结果

本文中所有分类实验的评价指标均由相同实验重复 5 次、计算平均值得到，从而避免实验结果的偶然性。作为对照组的随机森林（RF）方法运行环境、实验步骤均与 XGBoost 相同。在直接分类和二分类实验中，两种算法得到的结果混淆矩阵和分类报告如表 2~4 所示（NS 表示 notspam）。

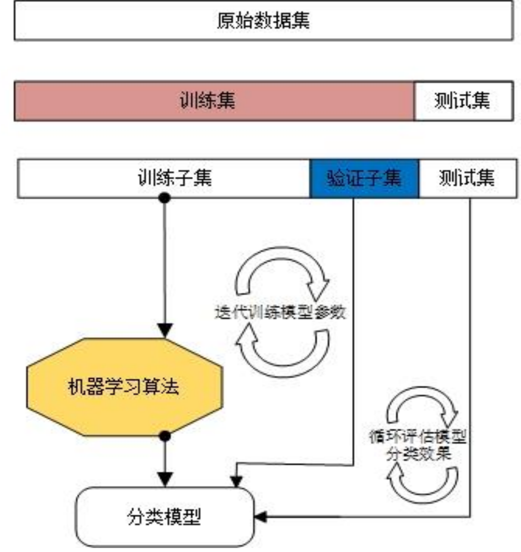


图 3　实验步骤展示图:内循环选择模型参数，外循环验证实验结果

Fig. 3　The steps of the experiment show: internal loop selecte model parameters, external loop validate test results

表 2　XGBoost 与 RF 二分类结果对比

Table 2　Comparison of classification results between XGBoost and RF

	准确率 P		召回率 R		F1 值	
数值（%）	XGBoost	RF	XGBoost	RF	XGBoost	RF
notspam	99.63	97.17	93.22	93.76	96.32	95.44
spam	41.60	78.15	<u>**93.22**</u>	<u>**89.10**</u>	57.53	83.27

表 3　XGBoost 与 RF 多分类结果对比表

Table 3　Comparison of XGBoost and RF results in multi-class classification

	准确率 P		召回率 R		F1 值	
数值（%）	XGBoost	RF	XGBoost	RF	XGBoost	RF
NS	87.66	85.08	94.32	95.56	90.88	90.02
BM	89.33	89.13	<u>**78.96**</u>	<u>**77.89**</u>	84.18	83.13
PL	69.61	70.58	<u>**68.66**</u>	<u>**64.74**</u>	69.24	67.54
BM	68.7	68.11	<u>**58.68**</u>	<u>**53.75**</u>	61.96	60.09

表 4　XGBoost 与 RF 实验结果混淆矩阵表（数值表示百分比）

Table 4　Confusion matrix of XGBoost and RF experimental results (value expressed percentage)

XGBoost 混淆矩阵					RF 混淆矩阵				
	BM	LM	NS	PL		BM	LM	NS	PL
BM	54.6	0.7	16.86	27.8	BM	54.2	1.4	17.2	27.2
LM	0.47	77.23	5.46	16.84	LM	1.4	77.6	5.3	16.6
NS	1.61	70.23	95.42	2.6	NS	1.4	0.5	95.7	2.4
PL	6.29	3.03	24.59	66.08	PL	7.4	4.1	23.3	65.2

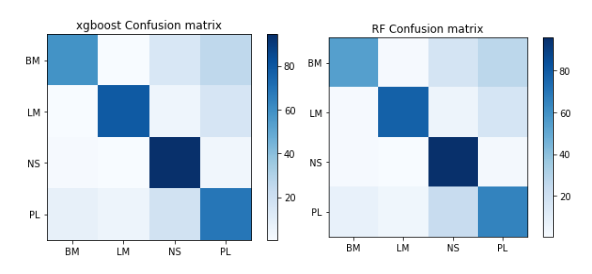


图 4 多分类混淆矩阵热力图

Fig. 4　Thermodynamic chart of multi class confusion matrix

由于在公安实战中，与正常用户相比，更关注异常用户识别率；与模型检测的准确率相比，更关注召回率，因此在对比算法性能时，应选取异常用户的召回率作为重要的评价指标。

表 3 及图 4 显示，混淆矩阵对角线所在色块颜色较深，说明两种方法均能实现对异常用户的有效监测。由以上表格可以看出，本文方法在将 spam 用户作为一个整体而进行的二分类任务中（表 5），对异常用户的总体检测召回率可以达到 93.22%，比随机森林的 89.11%提升了 4 个百分点。在多分类任务中（表 3）对各类 spam 用户的召回率分别达到 78.96%、68.66%、58.68%，与随机森林相比召回率以及 F1 值均产生稳定提升（召回率分别提升约 1%、4%、5%，F1 值提升 1%以上）。这表明本文方法对以检测异常用户为目标的公安实战工作更具实际意义。

3.3.2 不平衡数据集检测结果

本文通过保留全部正常用户、按比例随机剔除异常用户的方式构造不均衡数据集，分别构造了包括异常用户占全部用户比例为 10%-40%的数据集（异常用户比例 50%代表均衡数据集），但保持各数据集中三类异常用户的数量比例关系与原数据集相同。

表 5 为分别采用 XBGoost 和 RF 两种分类方式，在上述各数据集进行训练并测试的结果。表格中的数字为两种方法在对应的不平衡数据集中检测三类异常用户以及正常用户的准确率、召回率和 F_1 值。对比 XGBoost 和 RF 在表格相应位置的每个数据可以看出，两种集成方法在不均衡数据集中检测异常用户的能力甚至超过了均衡数据集，证明了集成学习方式处理不平衡数据的出色能力；并且 XGBoost 在对 BM 与 LM 两类异常用户检测的召回率显著高于 RF 方法，表明该方法对不平衡数据集更具有效性。

这是因为，XGBoost 在目标函数中二次项和正则项的共同作用下具有很强的泛化能力，在不平衡数据集中具有比 RF 更加优异的表现。此外，表格中 XGBoost 对不平衡数据集中 PL 用户检测的准确率略低于随机森林的原因可能是：PL 是数量最多的一种异常用户，在不平衡数据集中占比仍然较高；并且在收集和标注数据时，内容污染者（PL）是发布内容与话题标签信息不符的一类账号，与内容特征明显关联程度更高，保留全部特征的 XGBoost 方法对于保留部分特征的随机森林方法的优势不明显。

表 5 不均衡数据集中 XGBoost 与 RF 表现

Table 5　XGBoost and RF performance in imbalanced data sets

spam 用户比例	评价指标(%)	XGBoost				RF			
		BM	LM	NS	PL	BM	LM	NS	PL
50%	准确率	68.63	90.78	86.03	70.11	71.66	90.95	84.55	70.85
	召回率	56.16	78.01	95.23	66.41	52.96	76.35	96.51	65.42
	F_1 值	61.77	95.23	90.4	68.21	60.91	83.01	90.13	68.03
40%	准确率	74.38	89.61	99.98	76.75	74.71	90.39	99.92	74.2
	召回率	62.89	81.54	99.99	86.15	56.99	78.55	99.99	87.66
	F_1 值	68.15	85.38	99.99	81.18	64.66	84.05	99.96	80.37
30%	准确率	73.26	87.33	99.98	73.8	74.484	88.71	99.93	71.47
	召回率	60.52	78.98	99.99	84.25	55.81	76.15	99.99	86.1
	F_1 值	66.28	82.95	99.98	78.68	63.81	81.95	99.96	78.11
20%	准确率	73.29	84.57	99.99	72.45	76.02	86.93	99.93	70.74
	召回率	63.27	76.15	99.99	81.93	55	74.04	99.99	86.23
	F_1 值	67.91	80.14	99.99	76.9	63.82	79.97	99.97	77.56
10%	准确率	68.56	79.06	99.97	66.07	75.53	81.47	99.94	65.79
	召回率	54.38	72.48	99.99	76.57	47.87	71.04	99.99	83.31
	F_1 值	60.65	75.63	99.99	70.94	58.6	75.9	99.97	73.52

3.3.3 特征选择检测结果

社交网络用户的特征可分为文本、地点、用户、关系等四类。为探究不同类别的用户特征对分类结果的影响，并验证 xgboost 特征选择方法的有效性，本轮实验中分别选择四类特征单独训练模型，通过 XGBoost 按照影响力排名选择前 10 个、前 20 个特征单独训练 XGBoost 和 RF 分类器进行测试，做 5 次重复实验取平均值。分类效果如表 6 所示。

表 6　各类特征分类结果

Table 6　Classification results from different kinds of features

实验结果（%）		全部特征	前 10 个特征	前 20 个特征	文本特征	地点特征	用户特征	关系特征
平均召回率	XGBoost	82.9	73.2	81.25	73.83	68.48	58.54	59.64
	RF	81.2	72.85	80.68	73.05	66.56	58.21	59.54

实验表明，单独使用部分类别特征虽然也可以达到一定的分类效果，例如采用 32 个内容特征即可得到 73%的召回率，但通过 XGBoost 方法仅选择 20 个特征，就能在两种分类算法中实现 80%以上的平均召回率，接近采用全部特征的分类结果；仅采用前 10 个重要特征，仍能达到 73.3%的召回率，精度高于单独选取任何一类全部特征。这证明了社交网络异常用户检测过程中，综合选取各类特征可以达到比单独选取相同数量的某一类特征更为有效的结果，证明了 XGBoost 特征选择的有效性。在公安实战中，有效的特征选择过程可以减少样本采集所需的特征数，从而提升检测效率。此外，以上所有情况中 XGBoost 均得到比 RF 更高的召回率，再次证明 XGBoost 分类算法的优势。

4　结束语

社交网络异常用户检测本质上可以归结为分类或聚类问题。在构造决策树的过程中，XGBoost 算法在目标函数中对损失函数计算二次最优化，比只考虑一阶导数的梯度下降提升树等其他 boost 方法更具有全局搜索的能力，同时正则项的引入增加了模型的泛化性能，节点权重更新策略在保留特征完整信息的同时消除了非正交特征的影响，在社交网络 spam 用户的二分类、多分类检测任务中均获得出众效果。在更贴近社交网络实际情况的不平衡数据集进行 spam 用户检测时，XGBoost 表现更加优秀。在利用公开数据集识别 spam 的过程中，利用 XGBoost 进行特征选择，只保留三分之一的特征即可达到与选择所有特征相似的检测效果，可以提升数

据采集的效率。并且无论是选取全部特征或部分特征，XGBoost 与集成分类器随机森林方法相比，都得到召回率和 F1 值的提升，在公安工作中具有重要的实际意义。

研究展望：a）XGBoost 算法只能处理数值型特征，因此在实战检测过程中需要增加数据预处理的步骤，将非数字的特征进行数值型转换；b)不平衡数据中 XGBoost 和 RF 分别对不同类别的特征检测效果更好，因此可以根据不同的检测目标选择不同的分类方法；c)在今后的研究中可以将多种算法融合在异常用户检测模型中，从而提升社交网络异常用户检测模型的鲁棒性。

参考文献：

[1] Helen Costa, Luiz H. C. Merschmann, *et al.* Pollution, bad-mouthing, and local marketing: The underground of location-based social networks [J]. Information Sciences, 2014, 279: 123-137

[2] 张玉清，吕少卿，范丹．在线社交网络中异常账号检测方法研究 [J]. 计算机学报, 2015, 38(10): 2011-2027. (ZHANG Yuqing, LV Shaoqing, FAN Dan, Anomaly Detection in Online Social Networks [J] , Chinese Journal of Computers, 2015, 38(10): 2011-2027.)

[3] Miller Z, Dickinson B, Deitrick W, *et al.* Twitter spammer detection using data stream clustering [J]. Information Sciences, 2014, 260(1): 64-73.

[4] Henderson K, GallagherB, Li Lei, *et al.* It's who you know: graph mining using recursive structural features[C]//Proc of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York:ACM Press, 2011: 663-671.

[5] Chu Zi, Widjaja I, Wang Haining. Detecting social spam campaigns on Twitter [C]// Proc of International Conference on Applied Cryptography and Network Security. Berlin: Springer, 2012: 455-472.

[6] 郝亚洲,郑庆华，陈艳平,等.面向网络舆情数据的异常行为识别[J]. 计算机研究与发展,2016,53(3): 611-620. (Hao Yazhou, Zheng Qinghua, Chen Yanping, et al. Recognition of abnormal behavior based on data of public opinion on the Web [J]. Journal of Computer Research and Development, 2016, 53(3): 611-620.)

[7] Chen Tianqi,Tong He. Higgs boson discovery with boosted trees [C]//Proc of International Conference on High-Energy Physics and Machine Learning. 2014:69-80.

[8] Zheng Xianghan, Zeng Zhipeng, Chen Zheyi, *et al.* Detecting spammers on social networks[J]. Neurocomputing, 2015, 159(2): 27-34.

[9] 吕少卿．在线社交网络中异常账号检测研究[D]．西安:西安电子科技大学, 2016. (Lyu Shaoqing. Research on anomaly detection in online social networks [D]. Xi'an:Xidian University, 2016.)

[10] 刘琛．基于行为分析的社交网络异常账号的检测 [D].北京:北京交通大学, 2017. (Liu Chen. The detection of anomaly accounts based on behavioral analysis for social networks [D]. Beijing Jiaotong University, 2017.)

[11] Jiang Meng, Cui Peng, Beutel A,*et al.* CatchSync: catching synchronized behavior in large directed graphs[C]//Proc of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2014: 941-950.

[12] Yang Zhi, Xue Jilong, Yang Xiaoyong, *et al.* VoteTrust: leveraging friend invitation graph to defend against social network sybils [J]. IEEE Trans on Dependable & Secure Computing, 2016, 13(4): 488-501.

[13] Chen Tianqi, Carlos Guestrin. XGBoost: A scalable tree boosting system [C]//Proc of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press, 2016: 785-794.

[14] 张青青.非平衡类等异常检测研究[D].南京:南京航空航天大学, 2010. (Zhang Qingqing. Anomaly detection research fot imbalanced classes [D]. Nanjing:Nanjing University of Aeronautics and Astronautics, 2010.)

[15] Laurikkala J. Improving identification of difficult small classes by balancing class distribution[C]//Proc of Conference on AI in Medicine in Europe: Artificial Intelligence Medicine. Springer-Verlag,2001: 63-66.

[16] Yen S J, Lee Y S. Cluster-based under-sampling approaches for imbalanced data distributions [J]. Expert Systems with Applications, 2009, 36(3): 5718-5727.

[17] Chawla N V, Bowyer K W, Hal L O, *et al.* SMOTE: synthetic minority over-sampling technique [J]. Journal of Artificial Intelligence Research, 2002, 16(1): 321-357.

[18] Ertekin S, Huang J, Bottou L, *et al.* Learning on the border: active learning in imbalanced data classification[C]//Proc of the 16th ACM Conference on Conference on Information & Knowledge Management. New York:ACM Press, 2007: 127-136.

[19] 李克文,杨磊,刘文英,等.基于 RSBoost 算法的不平衡数据分类方法 [J]. 计算机科学, 2015, 42(9): 49-252. Li Kewen, Yang Lei, Liu Wenying, et al. Classification method of imbalanced data based on RSBoost [J]. Computer Science, 2015, 42(9): 249-252)